Using kittens to unlock photo-sharing website datasets

Simon Gascoin

Using kittens to unlock photo-sharing website datasets

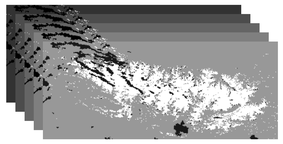
Simon Gascoin - CNRS/Centre d'Etudes Spatiales de la Biosphère, Toulouse, France. *simon.gascoin@cesbio.cnes.fr*

Context

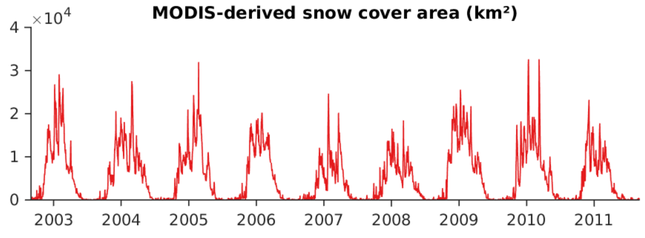
Mining photo-sharing websites is a promising approach to complement in situ and satellite observations of the environment, however a challenge is to deal with the large degree of noise inherent to online social datasets.

Data

The method was evaluated on a time series of the snow cover area in the Pyrenees.

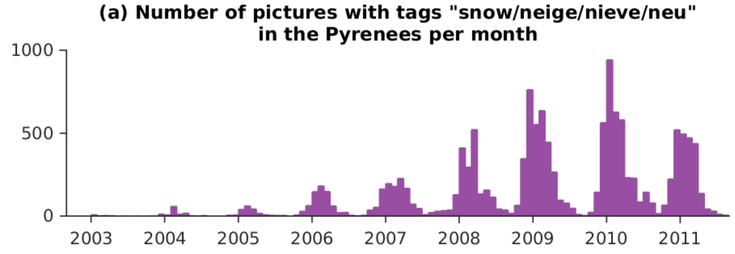


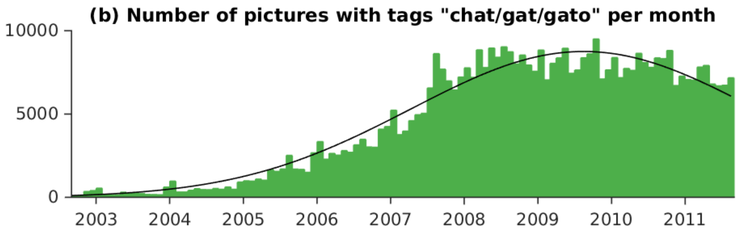
MODIS snow maps were processed to generate a daily cloud-free snow cover climatology over 2000-2013 (Gascoin et al., 2015).

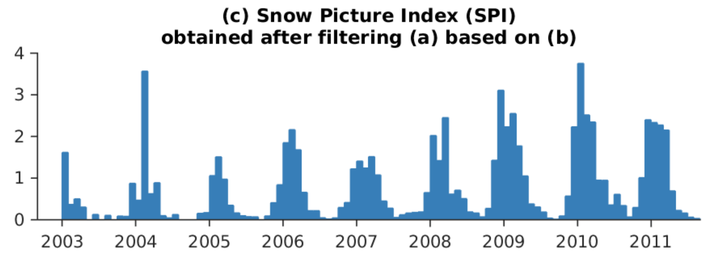


Method

Using the Flickr application programming interface all the public images metadata tagged at least with one of the following words were queried: “snow”, “neige”, “nieve”, “"neu” (snow in French, Spanish and Catalan languages). The search was limited to the geotagged pictures in the Pyrenees area. However, the number of public pictures available for a given time interval depends on several factors, including the Flickr website popularity and the development of digital photography.



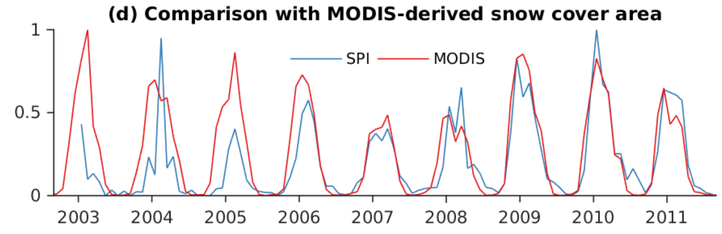




Then all images tagged with “chat”, “gat” or “gato” (cat in French, Spanish and Catalan languages) were queried. The tag “cat” was not considered in order to exclude the results from North America where Flickr got popular earlier than in Europe. The number of “cat” images per month was used to fit a gauss model of the number of images uploaded in Flickr with time. This model was used to remove this trend in the numbers of snow-tagged photographs.

Result

The comparison with MODIS snow cover area shows that the method effectively removes the first-order trend in the flick data (Spearman’s R increases from 0.5 to 0.8)



Conclusion

This study was restricted to 2003-2011. Since 2011 the advent of smartphones with built-in GPS and camera has strongly increased the amount of geotagged data. These data have the potential to overcome some limitations of remote sensing products like cloud obstruction provided that (i) the users choose a public sharing license (ii) the hosting website provides open source API to programmers.

References

Gascoin, S., Hagolle, O., Huc, M., Jarlan, L., Dejoux, J.-F., Szczypta, C., Marti, R., and Sánchez, R.: A snow cover climatology for the Pyrenees from MODIS snow products, Hydrol. Earth Syst. Sci., 19, 2337-2351, doi:10.5194/hess-19-2337-2015, 2015.

Gascoin, S. Using kittens to unlock photo-sharing website datasets. Journal of Brief Ideas, doi:10.5281/zenodo.44809, 2016.